

Copyright (c) 1993 - 2003 Compugen Ltd.

Gencore version 5.1.5

OM nucleic - nucleic search, using sw model

Run on: June 1, 2003, 15:21:24 ; Search time 2494.08 Seconds

(without alignments)
12850.800 Million cell updates/sec

Title: US-09-625-573-3
Perfect score: 19/9
Sequence: 1 CAGGACTGCCTGAGACAAGC.....ATATGCAATAAATTAG 1979

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:
1: em_estba:
2: em_estba:
3: em_estbaum:
4: em_estmu:
5: em_estov:
6: em_estpl:
7: em_estro:
8: em_ntcc:
9: qb_est1:
10: qb_est2:
11: qb_ntcc:
12: qb_est3:
13: qb_est4:
14: qb_est5:
15: em_estfun:
16: em_estom:
17: qb_gss:
18: em_gss_hum:
19: em_gss_inv:
20: em_gss_pnl:
21: em_gss_vrt:
22: em_gss_fun:
23: em_gss_mam:
24: em_gss_mus:
25: em_gss_other:
26: em_gss_pro:
27: em_gss_rdt:
1074 14 BM917063
14.6 14.4 452 9 AL851510
289 14.4 452 9 AL552677
c 8 284.8 14.4 452 9 AL552677
c 9 276.6 14.0 907 9 AL552677
c 10 249.8 12.6 789 12 BG205056
c 11 247.8 12.5 341 17 AQ066245
c 12 239.4 12.1 869 12 BF119225
c 13 236.4 11.9 876 13 BI906283
c 14 235.2 11.9 884 12 BG182330
c 15 230 11.6 230 12 BF117452
c 16 220.8 11.2 952 14 BM917763
c 17 216.8 11.0 1167 14 BQ053336
c 18 214.6 10.8 650 10 BB630138
c 19 198.6 10.0 650 10 BB630138
c 20 196.2 9.9 672 10 BB630138
c 21 196.2 9.9 568 12 BF193051
c 22 189 9.6 780 13 BI908127
c 23 186.8 9.4 551 12 BF081124
c 24 181 9.1 863 13 BI661279
c 25 180.6 9.1 669 13 BI393893
c 26 179.4 9.1 533 12 BF193051
c 27 177.2 9.0 639 10 BB629533
c 28 175.2 8.9 723 14 BM91933
c 29 173.6 8.8 965 14 BQ944555
c 30 172.2 8.7 934 12 BG460984
c 31 169.2 8.5 787 9 AU080004
c 32 168.8 8.5 3005 11 AK019478
c 33 163.8 8.3 936 12 BF119806
c 34 153.2 7.7 665 10 BB628567
c 35 151.8 7.7 662 10 BB644135
c 36 151.6 7.7 600 13 BI987229
c 37 151.6 7.7 646 10 BB615654
c 38 147.6 7.4 867 13 BI106339
c 39 146.4 7.4 2111 11 AK007808
c 40 145.4 7.3 319 14 BM734515
c 41 141.2 7.1 574 17 AQ461727
c 42 140.6 7.1 804 13 BI685810
c 43 139.6 7.1 427 12 BF081123
c 44 138.2 7.0 668 10 BB645326
c 45 138 7.0 770 9 AJ452673

ALIGNMENTS

RESULT 1
BI764263 LOCUS MGC_116
DEFINITION Homo sapiens mRNA sequence.
ORGANISM Homo sapiens
ACCESSION BI764263
VERSION BI764263.1
KEYWORDS EST
SOURCE Human
COMMENT GT:15755841
REFERENCE 1
AUTHORS 1
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
CONTACT Robert Strausberg, Ph.D.
Email: cospbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILNL)

DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the LILNL at:
<http://image.lilnl.gov>
Plate: LILNL1465 row: 1 column: 13
High quality sequence stop: 805.
Location/Qualifiers 1 . 807

SUMMARIES

Result No.	Score	Query	Match Length	DB	ID	Description
1	469.8	23.7	807	13	BT764263	BT764263 60304953
2	469	23.7	469	17	BB4215	BB4215 RPT111-22K6
3	357.6	18.1	487	10	BB656336	BB656336 UI-M-BH0-
4	326.2	16.5	542	9	AA547303	AA547303 vN28h04 r
5	296.4	15.0	480	10	AW657263	AW657263 109996 MA
6	292.6	14.8	447	9	AA671573	AA671573 v104g01.r

Db	241	GGGGAGGAGGAGTCTCGGTGGTTATAAAACGAGGAGGAGTTTCATTTGGTGTATTAA	300	
Db	1194	AGGGATAACAATCTGATATAAACAAACTTCAAGGGTTGTGAACATAAGAGAAC	1253	
Db	301	AGGGATAACAATCTGATATAAACAAACTTCAAGGGTTGTGAACATAAGAGAAC	360	
Db	1254	TGTAAGCAGGTGCCAGGAACCTCAGGGCTGTGTACTATACAGACTATGTACCCA	1313	
Db	361	TGTAAGCAGGTGCCAGGAACCTCAGGGCTGTGTACTATACAGACTATGTACCCA	420	
Qy	1314	ATGCATATCCAAACATGTGTCAGGGATAATAACGAAAAACTGTGGGTA	1362	
Db	421	ATGCATATCCAAACATGTGTCAGGGATAATAACGAAAAACTGTGGGTA	469	
RESULT 3				
LOCUS	BE656336	BE656336	487 bp	tRNA
DEFINITION	UI-M-BHO-aju-h-06-0-UI-1	NIH_BMAP_M_S1	Mus musculus	cDNA clone
ACCESSION	BE656336	UI-M-BHO-aju-h-06-0-UI-5'		mRNA sequence.
VERSION	BE656336.1	GT:99B2249		
KEYWORDS	EST.			
SOURCE	Mus musculus			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus .			
REFERENCE	1	(bases 1 to 487)		
AUTHORS	Bonaldo,M.F., Lennon,G. and Soares,M.B.			
TITLE	Normalization and subtraction: two approaches to facilitate gene discovery			
JOURNAL	Genome Res.	6 (9), 791-806 (1996)		
MEDLINE	9704477			
COMMENT	Contact: Chin, H National Institute of Mental Health 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD 20892-9643, USA Tel: 301 443 1706 Fax: 301 443 9890 Email: mestr@mail.nih.gov			
FEATURES	source			
ORIGIN				
BASE COUNT	115	a	103	g
ORIGIN				
Query Match	18.1%		Score 357.6;	DB 10;
Best Local Similarity	85.3%		Pred. No. 7e-89;	Length 487;
Matches	399		Mismatches 0;	Gaps 0;
FEATURES	source			
LOCATION/Qualifiers				
1.	.487			
/organism="Mus musculus"				
/strain="C57BL/6J"				
/db_xref="taxon:10090"				
/clone="UI-M-BHO-aju-h-06-0-UI"				
/dev_stage="N/H_32 days"				
/lab_host="DRI10B (Life Technologies)"				
/note="Vector: pRT3D-Pac (Pharmacia) with a modified polylinker; Site:1; Not I; Eco RI; The NIH_BMAP_M_S1 library is a subtracted library derived from a mixture of normalized libraries from ten regions of the mouse brain (cerebellum, brain stems, olfactory bulb, hypothalamus, cortex, amygdala, basal ganglia, pineal, gland, striatum, hippocampus). The driver used for subtraction consisted of a pool of 20,000 cDNA clones obtained from non-normalized and normalized libraries of these ten regions of the mouse brain."				
BASE COUNT	115	a	103	g
ORIGIN				
Query Match	18.1%		Score 357.6;	DB 10;
Best Local Similarity	85.3%		Pred. No. 7e-89;	Length 487;
Matches	399		Mismatches 0;	Gaps 0;

Page 8

/note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."/>

REFERENCE	1 (bases 1 to 869) NIH_MGC http://mgc.ncbi.nih.gov/.	DEFINITION	603063222F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5212541 5', mRNA sequence.
AUTHORS	National Institutes of Health, Mammalian Gene Collection (MGC)	ACCESSION	BI906283
TITLE	Unpublished (1999)	VERSION	BI906283.1
JOURNAL		KEYWORDS	GI:16168946
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgsaps@email.nih.gov Tissue Preparation: Lothar Hennighausen Ph.D., Robin Humphreys cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov/plate/LIAM9192 row: o column: 07 High quality sequence stop: 583.	SOURCE	Human
FEATURES	Location/Qualifiers 1..869 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone_id=NCI_GAP_Mam5" /clone_host=DH10B" /dev_stage="7 months" /lab_host="NIH" /note="Organ: mammary; Vector: PCMV-SPORT6; Site_1: Sali; Site_2: NotI; Cloned unidirectionally; Primer: Oligo dT. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Robin Humphreys, NIH"	source	
BASE COUNT	218 a 190 c 223 g 238 t	FEATURES	source
ORIGIN		source	
Query Match	12.1%; Score 239.4; DB 12; Length 869;	source	
Best Local Similarity	78.2%; Pred. No. 1..1e-55;	source	
Matches	351; Conservative 0; Mismatches 91; Indels 7; Gaps 5;	source	
QY	787 AGAAGAGCCATAGGGCAGTGAGAGTCATCTTCAACCATTGATGANTGTTCATTTCTCTCT 846	source	
Db	1 AGAAGAGGCACAGGGCTGAGGCTCATCTTCCATATGATTTCTCTCTCTCTCTCT 60	source	
QY	847 GGACTCCCTATAACATATGTCATTCATCTCCCTGAACACCTTCAGGAAATTCTGGGCTGACTA 906	source	
Db	61 GGACTCCATACAATA-GTTCCTCTTCATGACCACTTCAGGAACTCTGGGATGACTA 119	source	
QY	907 ACTGTGAAGCACCGTAACTTGACCAAGGCCAGGTGACAGACTCTGGGATGA 966	source	
Db	120 ACTGTGTGATTTGACAAAGACTTACGACGGCCATGACAGACTCTGGAAATGA 179	source	
QY	967 CTCACTGTGATCATATCCATATCATATGCTTCATGCTTGAGGTTCAGAGGTTC 1026	source	
Db	180 CAGACTGGCATTAATCTGTCAATTATGCTT-GRTGAGAAGACTTCGGAAAGGTATC 238	source	
QY	1027 TCTCGGTGTTCTTCGAAGACATCCAAAGGCTTCTGCAAACAATCTCCAGTTCT 1086	source	
Db	239 TCTCCATATTTCAGAAAGCACAT-GCTAAACGCTCTGCAAACAGTGGCCAGTTCT 297	source	
QY	1087 ACAGGGAGACAGTGATGATGAACTTACAAACACCGCTTCACCTGGGAGGGAG 1146	source	
Db	298 ATAGGGAGACAGGAGATCAGTGGCTTACATTCACCTTCACAGGGAAAGGG 357	source	
QY	1147 TCTCGGTGTTATAAACAGGAGGAGTGTGTTATAA--AGGGAGATAA 1203	source	
Db	358 TCTGG-TGGTTGTAAGTAAGTGACATTGGCTTTAAAGCAGGAAGGGAG 416	source	
QY	1204 CAATCTGTATAACAAACACTCAAGG 1232	source	
Db	417 CAGTTGTACATGCAACAGCTCAAAG 445	source	
RESULT	13	LOCUS	BI906283
		bp	876
		mRNA	linear
		EST	16-Oct-2001

Oncology Research Network
 University Health Network
 610 University Ave., 5-126, Toronto, Ontario, M5G 2M9, Canada
 Tel: (416) 946-4639
 Fax: (416) 946-6546
 Email: k.stewart@utoronto.ca
 PCR PRIMERS
 FORWARD: 5' -GCCAAGCTGAAATTACCCCTCACTAAAGGG-3'
 BACKWARD: 5' -CCAGTGATTGTATAAGCTACTATAGACATCACTATAGGGG-3'
 Seq primer: 5' GAANTAAACCCCTCACTAAAGGG-3'

FEATURES

source

1. .230
 /organism="Homo sapiens"
 /db_xref="Taxon:9606"
 /clone.lib=Myeloma (PCL) cDNA library"

/sex="male"
 /tissue_type="Blood"
 /cell_type="myeloma"
 /dev_stage="Plasma cell leukemia"
 /note="Vector: Lambda Zap Express; Site_1: EcoRI; Site_2:
 XbaI; mRNA was purified from plasma cell leukemia
 patient's peripheral blood containing >95% myeloma. An
 oligo d(T)18 primer containing XbaI restriction site was
 used to prime first strand synthesis using M-MuLV reverse
 transcriptase. To protect the cDNAs from XbaI digestion in
 subsequent cloning step, the nucleotide analogue
 5-methyl-dCTP was added to the nucleotide mixture and
 [a-32P]dATP was added to monitor the quantity and quality
 of first strand synthesis. After second strand synthesis
 and blunting of cDNA termini, EcoRI adaptors were ligated
 followed by kinase treatment and digestion with XbaI.
 The cDNAs were then size-fractionated on Sephadex G-
 500 column and then ligated into EcoRI and XbaI digested
 Lambda Zap Express vector. The ligation product was
 packaged using Gigapack II packaging extract. The library
 had primary titer of approx. 1x106. Clones from the
 primary library were randomly selected for single pass
 sequencing."

BASE COUNT 67 a 53 C 45 g 65 t
 ORIGIN

	Query	Match	Score	Length
Best	Matches	Local Similarity	11.6%	230;
Matches	230;	Conservative	100.0%	Pred. No. 3e-53;
				Mismatches 0;
				Indels 0;
				Gaps 0;
QY	1306	GTCACCCATGCAATGCAATATCCAACATGTCAGGAAATATCCAGAAA AATCTGCGGGAGAG 1365		
Db	1	GTCACCCATGCAATATCCAACATGTCAGGAAATATCCAGAAA AATCTGCGGGAGAG 60		
QY	1366	ACTTTGACTCTCCAGAAGCTCATCTAGCTCTGAAATGCCTCATTTGCTTA 1425		
Db	61	ACTTTGACTCTCCAGAAGCTCATCTAGCTCTGAAATGCCTCATTTGCTTA 120		
QY	1486	ATCCCTCTTCTCTAGTCATTAATTCTTCACTCAATCTGATTCTGCAATGTC TGTCTG 1485		
Db	121	ATCCCTCTTCTCTAGTCATTAATTCTTCACTCAATCTGATTCTGCAATGTC TGTCTG 180		
QY	1486	AAATCAAGGGCCAGCTGGAGGTGAAGAAGAGAATGTCACAGSCACAGATG 1535		
Db	181	AAATCAAGGGCCAGCTGGAGGTGAAGAAGAGAATGTCACAGSCACAGATG 230		

Search completed: June 1, 2003, 20:03:43
 Job time : 2502.08 secs



8